

1
AEROSOL PROCESS FOR FABRICATING DISCONTINUOUS
FLOATING GATE MICROELECTRONIC DEVICES

5
STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR
DEVELOPMENT

The U.S. Government has certain rights in this invention pursuant to grant DMR-9871850 awarded by the National Science Foundation.

10
RELATED APPLICATIONS

This application claims priority of U.S. provisional application serial No. 60/215,390, entitled AEROSOL PROCESS FOR FABRICATING DISCONTINUOUS FLOATING GATE MICROELECTRONIC DEVICES, filed on June 29, 2000, and U.S. provisional application serial No. 60/215,400, entitled DISCONTINUOUS FLOATING GATE INCORPORATING AEROSOL NANOPARTICLES, filed on June 29, 2000.

This application is related to U.S. application --/---,---, entitled AEROSOL SILICON NANOPARTICLES FOR USE IN SEMICONDUCTOR DEVICE FABRICATION, filed on June 29, 2001.

20
BACKGROUND OF THE INVENTION

Driven by a strong demand for portable electronic devices, non-volatile memory represents an important and rapidly growing sector of today's semiconductor memory market. Polysilicon floating gate memory devices conventionally have held the largest market share of non-volatile memory devices. In today's rapidly emerging non-volatile memory device sector, non-volatile floating gate memory devices have been fabricated by embedding silicon nanoparticles within the gate oxide of metal-oxide semiconductor field effect transistors (MOSFETs). Nanoparticles are so named because they include particle diameters on the nanometer scale. It is believed that silicon nanoparticle floating gate memory devices outperform conventional floating gate memory devices with faster read and write times, higher reliability, and lower power dissipation. The memory operation of nanoparticle field effect transistors depends on charge storage, similar to conventional non-volatile memory devices. In a nanoparticle non-volatile memory device, however, charge is not stored on a continuous floating gate polysilicon layer as in the conventional technology, but instead on a layer of discrete, preferably

1 crystalline silicon nanoparticles which may alternatively be referred to as nanocrystals
or quantum dots.

5 In these nanoparticle floating gate memory devices, the nanoparticles that act
as charge storage elements are located within the gate oxide of a MOSFET. Injecting
charge into the nanoparticles by tunneling from the channel alters the threshold voltage
of the transistor. A normal write/read/erase cycle includes information being written by
injecting charge from the channel into the nanoparticles, reading by measuring the
10 subthreshold current-voltage characteristics, and erasing by removing charge from the
particles to the channel. A single electron stored on each nanoparticle in an array with
a nanoparticle density of $3\text{-}10 \times 10^{11}/\text{cm}^2$ results in a threshold voltage shift of 0.3-0.5
volts that is easily detected at room temperature. Generally speaking, as compared to
conventional stacked-gate non-volatile memory devices, nanoparticle charge-storage
15 offers several potential advantages, such as: (1) simple, low cost device fabrication
since a dual-polysilicon process is not required; (2) superior retention characteristics
resulting from Coulomb blockade and quantum confinement effects, enabling the use
of thinner tunnel oxides and lower operating voltages; (3) improved anti-punchthrough
performance due to the absence of drain-to-floating gate coupling thereby reducing
drain induced punchthrough, allowing higher drain voltages during readout, shorter
20 channel lengths and consequently a smaller cell area; and (4) excellent immunity to
stress induced leakage current (SILC) and defects, due to the distributed nature of the
charge storage in the nanocrystal layer. Even if a significant fraction of the individual
nanocrystals that form the floating gate, are shorted to the channel/substrate, the non-
volatile memory device remains functional because the non-shorter nanocrystals
25 continue to store sufficient charge. The switching speed of devices made of
nanocrystal ensembles, however, is potentially limited by a distribution in charge transit
times, charging voltages, and threshold shifts resulting from various shortcomings of the
nanoparticle layer, such as the nanoparticle size and size distribution, nanoparticle
density, layer planarity and uniformity, and nanoparticle-to-nanoparticle interaction, i.e.,
30 lateral conduction.

Thus, there is a demonstrated need in the art for a layer of nanoparticles of
uniform size distribution and density. Similarly, there is a demonstrated need for
fabricating silicon or silicon-compatible nanocrystals with controlled size distributions
and oxide thicknesses that can be deposited on a substrate in a uniform and co-planar
35 manner. It is also desirable to fabricate the layer of nanocrystals using a process

1 sequence that is simple, reliable, low cost, easily controlled, repeatable, and free of
contamination. Previous attempts at producing a layer of nanocrystals suitable for use
in a field-effect transistor or other non-volatile memory devices, include the
5 shortcomings of uncontrolled particle sizes, non-uniformity of particle deposition, high
contamination levels, low density of the particle material, non-uniform density of the
particles within the nanoparticle layer, and unpredictable planarity of the nanoparticle
layer. Such irregular and unpredictable nanocrystal layers result in poor-performing or
non-functional devices.

10 In conclusion, in order to produce non-volatile memory devices with faster read
and write times, higher reliability and lower power dissipation, it is desirable to produce
nanocrystal floating-gate non-volatile memory devices using a simple, low cost
fabrication process which provides a layer of nanocrystals which forms a monolayer of
nanocrystals of uniform density and particle size.

15 SUMMARY OF THE INVENTION

To address these and other needs and in view of its purposes, the present
invention provides a process for forming a stratum of semiconductor or metal particles
having sizes in the nanometer range and suitable for application as the floating gate in
a non-volatile memory device. The stratum is composed of particles having a tightly
20 controlled range of particle sizes. The process includes decomposing a source of
semiconductor or metal material to form an aerosol of nanoparticles, then sintering or
heating the nanoparticles of the aerosol to convert the particles to crystalline material.
In an exemplary embodiment, the process may include quenching the nanoparticle
aerosol to minimize uncontrolled coagulation and to further control particle size. In an
25 exemplary embodiment, the majority of nanoparticles are single crystalline material.
The densified nanoparticles are compacted and include a density which approaches the
bulk density of the material of which they are formed. The densified nanoparticles may
optionally be classified by size, and particles outside the range of desired sizes are
removed from the aerosol stream.

30 The particles are preferably coated with a substantially continuous insulator
coating to produce particles having crystalline cores and a substantially continuous
insulating shell. The particles are then deposited onto a substrate surface using
thermophoretic or other means. The deposited, insulator-coated particles form a
35 stratum on the substrate surface, and in one embodiment, may be utilized as a floating

1 gate in a non-volatile memory device. The particles of the stratum are electrically isolated from one another.

5 BRIEF DESCRIPTION OF THE DRAWING

The invention is best understood from the following detailed description when read in conjunction with the accompanying drawing. It is emphasized that, according to common practice, the various features of the drawing are not to scale. On the contrary, the dimensions of the various features and the relative dimensions and locations of the features are arbitrarily expanded or reduced for clarity. Like numerals denote like features throughout the specification and drawing. Included in the drawing are the following figures.

Figure 1 is a flow chart of an exemplary process sequence of the present invention;

Figure 2 is a schematic of an exemplary processing system of the present invention;

Figure 3 is a cross-sectional view of an exemplary pyrolysis furnace of the present invention;

Figure 4 is a cross-sectional view of an exemplary radial differential mobility analyzer used to classify nanoparticles in the present invention;

Figure 5 is a cross-sectional view of an exemplary thermophoretic deposition system of the present invention;

Figure 6 is a cross-sectional view of an exemplary semiconductor nanoparticle of the present invention;

Figure 7 is a perspective view showing a stratum of semiconductor nanoparticles formed on a surface according to the present invention;

Figure 8 is a cross-sectional view of a stratum of semiconductor nanoparticles formed according to the present invention;

Figure 9 is a cross-sectional view showing a stratum of semiconductor nanoparticles of the present invention interposed between exemplary oxide layers;

Figure 10 is a cross-sectional view showing the gate structure of an exemplary non-volatile floating gate field effect transistor according to the present invention;

Figure 11 is a cross-sectional view of an exemplary non-volatile floating gate field effect transistor formed according to the present invention;

1 Figure 12 is a graph showing job sub-threshold characteristics of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention;

5 Figure 13 is a graph showing output characteristics of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention;

10 Figure 14 is a graph showing programming transients of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention;

15 Figure 15 is a graph showing erase transients of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention;

20 Figure 16 is a graph showing endurance characteristics of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention; and

25 Figure 17 is a graph showing retention and disturb characteristics of an exemplary MOSFET formed to include a floating gate formed of the exemplary silicon nanocrystals of the present invention.

30 DETAILED DESCRIPTION OF THE INVENTION

35 The present invention provides an aerosol process to form semiconductor nanocrystals entrained in a gas, and to deposit them on a substrate. The process sequence includes pyrolyzing or otherwise decomposing a semiconductor material or precursor to form semiconductor nanoparticles and controlling processing conditions to prevent uncontrolled agglomeration of the nanoparticles. The nanoparticles formed as an aerosol, may preferably be sintered to convert the particles to dense crystalline material, again using processing conditions chosen to allow for a controlled rate of agglomeration and so as to preclude the formation of undesirable agglomerates of large particles. The sintering process densifies the particles such that the cores are characterized by a density which approaches the bulk density of the material of which they are formed. An aspect of the present invention is the controlled size distribution of the produced crystalline nanoparticles, the densified nature of the nanoparticles, and that the majority of nanoparticles include a single-crystalline structure. The nanoparticles within the aerosol may advantageously be classified according to size,

1 such that the distribution of particle sizes is tightened further. Nanoparticles which
include sizes outside of a desired range of particle sizes are removed from the aerosol
prior to further processing. After classification, a dielectric shell is formed on the
5 nanoparticles and preferably encapsulates the cores. Thermal oxidation or other
methods may be used. The nanoparticles are then deposited on a substrate surface,
such as over the tunnel oxide of a non-volatile memory field effect transistor. The
deposited, dielectric-coated individual particles are electrically isolated from one
another.

10 The particles may be deposited as a monolayer and they may be deposited to
include a particle density ranging from 10^{12} to 10^{13} particles/centimeter. They may be
deposited such that a significant portion of adjacent particles contact each other. The
dielectric shells prevent lateral shorting between such contacting particles. The
deposition process may involve the formation of a colloidal suspension of the particles,
15 or thermophoretically or electrophoretically depositing the particles directly onto a
substrate. According to an exemplary embodiment, after the particles are deposited on
the substrate, they may be manipulated using force manipulation techniques in an
atomic force microscope or other equipment. In other exemplary embodiments,
colloidal forces and/or self-assembly techniques may be used to manipulate the
deposited particles. The manipulated particles may be arranged to form an ordered
20 structure such as a wire.

25 Another aspect of this invention is a process to produce a stratum of metallic
nanoparticles consisting of metallic cores and dielectric coatings. The metallic
nanoparticles are substantially similar to the semiconductor nanoparticles described
above, except that the core materials are metal. The process for forming the metal
nanoparticles and depositing the metal nanoparticles on a substrate to form a stratum
involves decomposing a metallic material or precursor to form metallic nanoparticles
entrained in a gas as an aerosol, then treating and depositing the particles of the
aerosol as described above in conjunction with semiconductor nanoparticles.

30 The stratum of semiconductor nanocrystals or metal nanoparticles produced by
the present invention, is suitable for use as the floating gate of a floating gate transistor
as well as for other applications in various semiconductor and microelectronic devices.
The stratum may be characterized as a discontinuous floating gate when used as the
floating gate of a transistor, and finds particular application in non-volatile memory
35 devices.

1 Now turning to the figures, Figure 1 shows an exemplary flow chart of the
process of the present invention. At step 400, an aerosol of semiconductor or metal
nanoparticles is formed by pyrolysis or other decomposition methods. At step 402, the
5 nanoparticles contained within the aerosol are sintered to form crystalline material. At
optional step 404, the crystalized nanoparticles are classified by size. Nanoparticles
which fall outside of a pre-selected range of particle sizes, are removed at 405B. The
in-range nanoparticles are next delivered for dielectric coating at 405A. At step 406, the
nanoparticles are oxidized to form an oxide shell over the nanoparticles, and at step
10 408 the nanoparticles of the aerosol are deposited on the substrate.

For brevity, the following detailed description will refer to the nanoparticles as
semiconductor nanoparticles, but it should be understood that the following description
applies to metal nanoparticles as well. With respect to embodiments specifically
applicable to metal nanoparticle formation and constitution, such will be pointed out.

15 Figure 2 shows a schematic of an exemplary arrangement of processing
equipment used to carry out the generalized process flow shown in Figure 1. The
arrangement shown in Figure 2 is exemplary only as other and additional equipment
may be used to carry out the process according to the present invention. In the
exemplary embodiment shown in Figure 2, pyrolysis furnace 14, differential mobility
20 analyzer 22 and oxidation furnace 32 are shown to be positioned outside of clean room
area 44 while deposition chamber 36 is shown to be positioned within clean room area
44. Such arrangement is exemplary only and in alternative embodiments all of the
equipment may be included within a clean room environment. In the first step of the
exemplary processing operation, a semiconductor material or precursor is decomposed
25 or broken down to form smaller units of semiconductor particles. Various methods for
decomposition of the semiconductor material may be used, including gas evaporation,
laser ablation, spark ablation, pyrolysis and various other chemical reactions.
According to the alternative embodiment in which a metallic material or precursor is
used to form metallic nanoparticles, gas evaporation, laser ablation and spark ablation
30 are advantageously utilized. Returning to the exemplary semiconductor embodiment,
the decomposition of the semiconductor material produces a plurality of discrete
semiconductor nanoparticles, so named because they typically include particle sizes
in the nanometer range. The semiconductor material from which the nanoparticles are
made by decomposition, may preferably be silicon but other semiconductor materials
35 and precursors and also other metallic materials and precursors may be used according

1 to alternative embodiments. Materials other than silicon are preferably selected for
compatibility with silicon processing. In an exemplary embodiment, a pyrolysis furnace
is used to form the plurality of discrete nanoparticles.

5 Referring to Figure 2, pyrolysis furnace 14 is fed by particle source gas 2 and
diluent gas 4. In an exemplary embodiment, diluent gas source 12 may be nitrogen but
other inert species such as helium may be used according to alternative embodiments.
In the exemplary embodiment shown, diluent gas 4 is filtered by filter 8 and the diluent
10 gas flow rate is controlled by MFC (mass flow controller) 18. Particle source gas 2
includes the source of semiconductor particles, preferably silicon, and includes a flow
rate controlled by valve 6. Particle source gas 2 originates from source 10 and will
preferably be a mixture also including a carrier gas. The arrangement shown in Figure
2 is exemplary only and it should be understood that other gas flow control means may
be used. For example, the flow rate of particle source gas 2 may preferably be
15 controlled by an MFC (not shown). In an exemplary embodiment, each of diluent gas
4 and the carrier gas may be nitrogen provided by a single source. The pyrolysis
furnace 14 is shown in greater detail in Figure 3.

20 Now turning to Figure 3, particle source gas 2 and diluent gas 4 are each
delivered to pyrolysis furnace 14. Particle source gas 2 includes a source from which
the semiconductor nanoparticles will be formed. In an exemplary embodiment, particle
source gas 2 may include silicon. In an exemplary embodiment, the silicon may be
delivered as silane in particle source gas 2. Gaseous silane may be considered a
silicon precursor. According to other exemplary embodiments, disilane or other silicon
sources may be used alternatively. In an exemplary embodiment, the carrier gas
25 included within particle source gas 2 may be nitrogen, but other inert carrier species
may also be used. Diluent gas 4 will be nitrogen in the preferred embodiment, but other
inert gases may be used alternatively as the diluent gas. According to one exemplary
embodiment, the carrier gas and diluent gas may emanate from the same source. In
a preferred embodiment, each of the diluent gas and carrier species may be nitrogen
30 and may be ultrahigh purity (UHP) nitrogen which flows through an oxygen getter, then
splits into two flows. One of the flows becomes diluent gas 4 and the other flow serves
as the carrier gas and flows along with the semiconductor material in particle source
gas 2. In an exemplary embodiment, each of diluent gas 4 and particle source gas 2
may include a flow rate of 750 sccm. According to other exemplary embodiments, the
35 flow rates for each of diluent gas 4 and particle source gas 2 may range from 650-850

1 sccm, but other flow rates may be used alternatively. According to a preferred
embodiment, particle source gas 2 includes 750 sccm of purified nitrogen as a carrier
gas in combination with less than 1 sccm of 5000 ppm silane in nitrogen. According to
5 other exemplary embodiments, other dilute mixtures of silane or other dilute mixtures
of various other semiconductor source gasses, may be used as particle source gas 2.
Pyrolysis furnace 14 includes maximum temperature zone 58 and pyrolysis zone 56.
Pyrolysis furnace 14 is heated using conventional heating means 52. The gas flows
enter pyrolysis furnace 14 in the entry or pyrolysis region 56 and flow downstream into
10 the maximum temperature zone 58. The flows of each of particle source gas 2 and
diluent gas 4 may be delivered to pyrolysis furnace 14 through 1/8" stainless steel
tubing, but other tubing may be used according to other exemplary embodiments.
Within pyrolysis furnace 14, stainless steel tubing may be used or quartz tubing may be
used. In a preferred embodiment, particle source gas 2 is directed to pyrolysis furnace
14 through inner tube 54 which is axially centered within outer tube 50 through which
15 diluent gas 4 enters pyrolysis furnace 14. In a preferred embodiment, the maximum
temperature zone of pyrolysis furnace 14 may be maintained at a temperature within
the range of 950°C to 1150°C, and at 1050°C in the preferred embodiment. Other
maximum temperatures may be used alternatively.

20 In the exemplary embodiment discussed herein, silane gas is the particle source
and silicon nanoparticles will be formed. It should be understood, however, that such
is exemplary and that other source gases, including other semiconductor or metal
materials or precursors, may be used to form other semiconductor or metal
nanoparticles according to other exemplary embodiments. Other exemplary source
25 gasses may include GaAs, GaN or p-doped semiconductor precursors.

Silane flow is introduced to pyrolysis furnace 14 as part of particle source gas 2
and flows through pyrolysis zone 56 towards maximum temperature zone 58. The gas
is heated in pyrolysis zone 56 during a ramp-up period during which it is heated to the
maximum temperature achieved in maximum temperature zone 58. The silane
30 pyrolyzes in the pyrolysis zone and breaks up into a plurality of discrete silicon
nanoparticles which form an aerosol. Stated alternatively, the plurality of discrete
semiconductor nanoparticles are entrained in a gas. The nanoparticles initially form by
homogeneous nucleation and grow by coagulation and vapor deposition. In an
exemplary embodiment, particle growth occurs during the ramp-up period during which
35 the gas is heated to the maximum temperature it achieves in the maximum temperature

1 zone. In an exemplary embodiment, particle growth occurs during a ramp-up period
ranging from 10-50 milliseconds, preferably 30 milliseconds. Gas flow rates are chosen
in conjunction with tubing size and length of the pyrolysis zone to achieve a residence
5 time of chosen duration in the pyrolysis zone. During the ramp-up time within pyrolysis
zone 56 of pyrolysis furnace 14, the silane pyrolyzes to form an aerosol of silicon
nanoparticles. Also during the ramp-up time within pyrolysis zone 56 and before
particle source gas 2 reaches maximum temperature zone 58, a diluent stream
preferably of UHP nitrogen combines in an annular flow around the aerosol particle
10 flow, upstream from the maximum temperature zone, to reduce particle concentration
and suppress particle coagulation. An exemplary arrangement shown in Figure 3
includes diluent gas 4 entering through outer tube 50 to quench particle source gas 2
which enters through inner tube 54, but other arrangements for mixing the respective
gasses may be used alternatively. Silicon aerosol particles with diameters ranging from
15 3-50 nanometers are produced by controlling the flow rate of the dilute silane and the
flow rates of the quenching ultrahigh purity nitrogen or other gases. In a preferred
embodiment, the diameters of each of the particles may be less than 20 nanometers.
The annular quenching flow also suppresses particle deposition onto the tube walls and
therefore reduces contamination. The flow of the gas mixture of the diluent gas and
20 particle source gas 2 continues in the maximum temperature zone 58.

25 Within maximum temperature zone 58, the silicon nanoparticles are sintered so
that they preferably become dense, spherical, single crystal nanoparticles. According
to another exemplary embodiment, the densified nanoparticles may become
polycrystalline particles. Flow conditions are chosen such that the sintering time may
range from 200-600 milliseconds, preferably 400 milliseconds. During this
densification/sintering process, it is essential that the reactor environment be oxygen-
free since oxygen will cause the formation of silicon oxides rather than crystalline
silicon. Additionally, oxygen prevents the desirable sintering of silicon nanoparticles.
The sintering is carried out in a controlled fashion to suppress uncontrolled coagulation
30 of the particles and therefore to maintain the size distribution of the particles within a
tight range. During this sintering process, the particles are heated to a sufficiently high
temperature to density the particles to form highly compacted particles, and to
crystallize the particles. The densified, discrete particles are formed to have a density
which is substantially as great as the bulk density of the pure semiconductor material
35 of which they are formed. The bulk density of a material is the maximum density of the

1 material at standard temperature and pressure (STP) conditions. Moreover, a majority
of the densified particles will be single crystalline material. In a preferred embodiment,
all of the densified particles will be single crystalline material. Also in the preferred
5 embodiment, the densified particles will be generally spherical in shape. Aerosol
stream 16, which includes the silicon nanocrystals, exits pyrolysis furnace 14.
According to another exemplary embodiment in which single crystalline particles are not
needed, the sintering process may be bypassed. After the sintering process, particle
diameters may range from 3-50 nanometers, preferably less than 20 nanometers.

10 Once formed and sintered, the densified silicon nanoparticles or nanocrystals of
the aerosol may advantageously be charged and classified by a size. The classification
may be accomplished by time-of-flight separations of a focused particle beam. Figure
2 shows aerosol stream 16 being delivered to an exemplary differential mobility
analyzer (DMA) 22 used to classify the nanoparticles. Aerosol stream 16 may optionally
15 be cooled prior to classification, using conventional means. Within DMA 22, the
nanoparticles within the aerosol are separated and out-of-range nanoparticles are
removed from the DMA 22 in out-of-range nanoparticle stream 24, while classified
particles within a pre-selected range of sizes continue to flow in stream 26 for additional
processing. DMA 22 is shown in greater detail in Figure 4.

20 To positively charge the nanoparticles for classification, they may be passed
through a so-called aerosol neutralizer prior to delivery to DMA 22. In the aerosol
neutralizer, the particles are exposed to an ambipolar cloud of gas ions produced by a
Kr source, in an exemplary embodiment. During this process, a small fraction of the
nanoparticles become charged with most carrying a single charge. Other devices for
25 charging the nanoparticles may be used for classification, particularly ones that change
a larger fraction of the particles.

Now turning to Figure 4, radial DMA 22 is a parallel-plate capacitor with radial
symmetry in an exemplary embodiment. According to other exemplary embodiments,
DMA 22 may take on other configurations. Generally speaking, an electric field is
30 applied to the plates of the differential mobility analyzer 22 to separate the charged
nanoparticles and to allow for particles having diameters within a pre-selected range of
particle sizes, to exit DMA 22 and continue processing. In the exemplary embodiment
shown, an electric field is applied across electrodes 62 and 64. Particle separation
occurs due to electromigration of particles within DMA 22. The desired, pre-selected
35 particle range may vary according to various exemplary embodiments. Positively

1 charged nanoparticles within aerosol stream 16 enter DMA 22 tangentially, through
aerosol inlet port 25 and migrate radially inward along a particle-free laminar sheath
flow 66 in the presence of an applied negative electric field. Ideally, the nanoparticles
5 carry only one elementary charge, and since these entering particles are generally
spherical, their mobility is related to their diameter. Nanoparticles with different sizes
follow different trajectories 68 within DMA 22. At the center of DMA 22 is sampling port
27 through which the classified nanoparticles within the pre-selected size range, are
removed. Uncharged nanoparticles and nanoparticles with sizes larger than the
10 classified size exit DMA 22 with an excess flow as part of exhaust flow 24. Particles
which are smaller than the classified particles in the pre-selected range deposit on the
bottom electrode of DMA 22. The diameter of the nanoparticles which are transmitted
in classified sample flow 26 is controlled and determined by flow rates and the applied
voltage.

15 The particle concentration within the aerosol stream can be monitored
throughout the synthesis process by flowing a portion of the classified aerosol flow into
an electrometer with femtoampere sensitivity. By varying the voltage applied to DMA
22 and detecting the particle concentration from the electrometer read downstream
from DMA 22, for example, the particle size of the silicon nanoparticles exiting the
furnace can be measured.

20 In an exemplary embodiment, particle diameters may be chosen to lie within the
range of 2-10 nanometers, but other ranges such as 2-6 nanometers, 3-7 nanometers,
5-10 nanometers, or less than 15 nanometers, may be chosen alternatively. In another
exemplary embodiment, nanoparticles within the 1 to 10 nanometer diameter size range
may be classified within about 6% on diameter. According to another exemplary
25 embodiment, the nanoparticles may be classified such that 90% of the particles are
within the desired range.

30 In cases where a thermal oxide passivating layer is desired, classified particle
flow 26 is next delivered to oxidation furnace 32 along with oxygen stream 29 as shown
in Figure 2. Oxygen source 28 and therefore oxygen flow 29 may be controlled by MFC
30 and include an oxygen concentration of 20% oxygen in nitrogen in an exemplary
embodiment. Oxidation may occur as the silicon nanoparticle aerosol is combined with
a controlled flow of oxygen or the silicon nanoparticle aerosol may be flowed through
a water evaporator. According to yet another exemplary embodiment, oxidation may
35 take place using TEOS (tetraethyl orthosilicate). According to still other exemplary

embodiments, dielectric or insulating layers other than an oxide may be formed to coat the individual nanoparticles to insure that the individual nanoparticles are electrically isolated from one another. The insulating-layer coated particles are preferably encapsulated by the insulating-layer coating.

In the oxidation furnace/oxygen flow embodiment shown in Figure 2, since the silicon nanoparticles are produced in ultra high purity nitrogen, the oxygen flow rate can be changed to alter the concentration of oxygen in the final oxygen/nitrogen mixture. Various oxygen flow rates may be used. The oxidation temperature within oxidation furnace 32 can also be varied. Typical oxidation furnace temperatures lie in the range of 700-1100°C with residence times of 200-800 milliseconds. Oxidation temperatures within oxidation furnace 32 are chosen in this range because lower temperatures produce poorer quality oxides, whereas higher temperatures have the tendency to evaporate silicon dioxide that is formed on the surface of the silicon nanoparticles. During the oxidation process which may occur at 900°C in an exemplary embodiment, a thermal oxide shell forms on the exposed silicon surface of the silicon nanoparticles, partially consuming the outer silicon material during the reaction. The thickness of the oxide shell is controlled by controlling the residence time of the aerosol within oxidation furnace 32. In general, a longer residence time increases oxide shell thickness. Oxide shell thicknesses typically range from 1-10 nanometers, and in a preferred embodiment may be controlled within the range of 1.5-2.0 nanometers. The thermal oxidation process is a uniform process which consumes the silicon surface of the preferably crystallized silicon nanoparticles thereby reducing the "core" size of the silicon nanoparticle. After oxidation, the silicon nanoparticles preferably include a crystalline silicon core having a diameter uniformly reduced with respect to the diameter of the silicon particle prior to oxidation. The silicon nanoparticle also includes an oxide shell surrounding the silicon core. This will be shown in Figure 6.

According to another exemplary embodiment, other materials such as water can be used to oxidize the silicon nanoparticles in a hot tube furnace, a process known as steam oxidation. According to this embodiment, the aerosol of silicon nanoparticles flows through a water evaporator (not shown) maintained at a controlled but variable temperature to control the amount of water allowed into the oxidation furnace. The water concentration can be controlled by varying the aerosol flow rate through the water evaporator. After the water evaporator, the aerosol and water are flowed through a high

1 temperature oxidation furnace where the particles become oxide passivated at
temperatures of 700°C-1000°C.

5 According to yet another exemplary embodiment, the silicon nanoparticles may
be coated with silica by chemical vapor deposition using TEOS (tetraethyl orthosilicate)
or other SiO₂ precursors such as TMOS (tetramethyl orthosilicate). Using the TEOS
oxidation process, the particle core size is not reduced as silicon material is not
10 consumed during the deposition of a silicon dioxide film onto the classified silicon
nanoparticles using TEOS as the precursor. The deposition process allows the
possibility of producing oxide-passivated silicon nanoparticles while retaining the silicon
core diameter such that it will not differ significantly from the original classified core
diameter. To produce silicon nanoparticles which are uniformly coated with a silicon
dioxide film, temperatures below about 750°C are used. In an exemplary embodiment,
15 using oxidation temperatures ranging from 500°C to 700°C, an oxide thickness of about
0.6nm is uniformly deposited on the particles. Under certain circumstances, the TEOS
does not form a uniform layer around the particle, rather, the TEOS does not wet the
silicon surface and, thus, balls up on the particle surface. When introduced into the
oxidation furnace, the TEOS ball then reacts to produce silicon dioxide and the final
20 particle morphology is the originally classified silicon nanoparticle with a silicon dioxide
particle attached to its surface. To prevent this undesirable particle morphology, it is
necessary to pre-treat the silicon nanoparticle surface. One pre-treatment method is
to expose the nanoparticle aerosol to ethanol vapor but other materials such as
methanol or isopropanol may be used in other exemplary embodiments. This involves
25 flowing the aerosol through an ethanol evaporator maintained at a constant temperature
to control the amount of ethanol introduced into the aerosol flow. This pre-treatment
technique has been found to be particularly advantageous for TEOS oxide passivation
performed at temperatures below about 500°C.

30 According to other exemplary embodiments, other chemical vapor deposition
methods and other methods for causing the surface of the nanoparticles to react with
a gas may be used. Various dielectric or insulating layers may be formed substantially
continuously over the nanoparticle surface, thereby forming a dielectric shell of sorts.

35 After the silicon nanocrystals have preferably been coated with a dielectric
coating according to one of the exemplary methods, the particles are collected for
subsequent use in device fabrication. According to one exemplary embodiment, the
particles may be directly deposited onto a device substrate and, according to another

1 exemplary embodiment, the particles may be collected in liquid for subsequent colloidal
processing. The former approach has the advantage of minimizing potential
contamination and ease of interfacing with the scanning probe microscope and other
5 dry manipulation techniques. The latter approach offers the potential to use colloidal
forces to assist in the assembly of the classified, passivated silicon nanoparticles into
the desired device structures. Several techniques for forming two-dimensional
monolayer arrays of silicon nanoparticles from stabilized colloids are available. In an
exemplary embodiment, electrophoretic or electrostatic deposition may be used.
10 According to electrophoretic deposition, charged particles are caused to migrate by
application of an electric field and to be deposited on a surface.

According to one exemplary embodiment, conventional electrophoretic
deposition techniques may be used to deposit the exemplary oxidized nanoparticles
directly onto a substrate surface. According to another exemplary embodiment,
15 deposition may be accomplished by inertial impaction. In yet another exemplary
embodiment such as shown in Figure 2, thermophoretic deposition techniques may be
used to deposit the silicon nanoparticles from the oxidized aerosol directly onto a
substrate surface. Figure 2 shows exemplary oxidized particle stream 34 being
delivered to thermophoretic deposition chamber 36. Thermophoretic deposition
20 chamber 36 includes water inlet 38 and water outlet 40 to cool the substrate within the
thermophoretic deposition chamber. Thermophoretic deposition chamber 36 is shown
in more detail in Figure 5.

Now turning to Figure 5, thermophoretic deposition chamber 36 includes aerosol
inlet port 35 through which oxidized particle stream 34 enters thermophoretic deposition
chamber 36. In the exemplary deposition chamber shown in Figure 5, aerosol oxidized
25 particle stream 34 enters the chamber centrally and moves radially outward through the
chamber and exits as gas outlet stream 84. Substrate 74 including substrate surface
76 rests on water-cooled disk 86. Water-cooled disk 86 is cooled by cooling water
which enters as inlet water 38 and exits as outlet water 40 and is preferably maintained
at a temperature of 23°C, although other temperatures may be used according to other
30 exemplary embodiments. Moreover, various other suitable methods may be used to
cool the underside of substrate 74 and to maintain it at a significantly lower temperature
than other portions of the deposition chamber. The top of the deposition chamber is
heated to 200°C using conventional means in an exemplary embodiment, although
35 other temperatures may be used according to other exemplary embodiments. This

1 provides an effective temperature difference of 177°C. According to various other
exemplary embodiments, temperature differences in the deposition chamber may be
about 150°C, 175°C, 200°C, or values therebetween. In a preferred embodiment, the
5 temperature difference between portions of the preferably metal deposition chamber
and the water-cooled disk 86 may be at least 170°C. Since the nanoparticles within the
aerosol move from high temperatures to cooler temperatures, the particles deposit on
surface 76 of substrate 74. To preclude any contact between the deposition chamber
and the substrate 74, substrate 74 may be supported by quartz pegs 82 in an
10 exemplary embodiment. According to other exemplary embodiments, quartz pegs 82
may not be needed. In a preferred embodiment, substrate 74 may be a silicon wafer
commonly used in the semiconductor manufacturing industry and surface 76 may
include the substructure of a device onto which a monolayer of silicon nanocrystals is
preferably deposited. According to other exemplary embodiments, other substrates
15 upon which semiconductor or microelectronic devices will be formed, may be used.
Thermophoretic deposition chamber 36 may be sized to accommodate various sizes
of substrates 74 such as a 6" silicon wafer, an 8" silicon wafer, and other substrates of
various sizes.

The thermophoretically deposited layer of silicon nanoparticles may be
20 characterized as a stratum composed of a monolayer of nanoparticles or simply as a
discontinuous layer of silicon. In an exemplary embodiment, particle densities may
range from 10^{12} to 10^{13} particles per square centimeter.

Figures 6-10 show various characteristics of the stratum of oxidized discrete
silicon nanoparticles. Figure 6 is a cross-sectional view showing a single nanoparticle
100. Nanoparticle 100 includes core 102 and oxide shell 106. Each of core 102 and
nanoparticle 100 are preferably spherical in shape. In an exemplary embodiment,
25 diameter 104 of core 102 may be less than 15 nanometers and may range from 2-10
nanometers, most preferably 3-7 nanometers, according to various exemplary
embodiments. In an exemplary embodiment, at least 90% of the cores will include a
diameter between 2 and 10 nanometers. Dielectric shell 106 may include a thickness
ranging from 1-10 nanometers and will include a uniform thickness of 1.5 to 2.0
30 nanometers in an exemplary embodiment. In a preferred embodiment, dielectric shell
106 will be an oxide shell. Dielectric shell 106 is a continuous dielectric coating which
envelopes core 102. In an exemplary embodiment, silicon core 102 may represent 30-
35 75% of the volume of silicon nanoparticle 100, including oxide shell 106. Core 102 is

1 formed of a semiconductor material compatible with silicon processing. In the preferred
embodiment, core 102 is formed of silicon, and will preferably be a crystalline silicon
material. In a preferred embodiment, core 102 will be a single crystalline material.
5 Core 102 includes a density at or near the bulk density of the pure semiconductor
material of which it is formed. The bulk density of a material is the maximum density
of the material at standard temperature and pressure (STP) conditions.

Now turning to Figure 7, a perspective view of stratum 101 of a plurality of
semiconductor nanoparticles 100 is shown as being formed on substrate surface 76.
10 The nanoparticles which make up stratum 101 are as described in Figure 6. An aspect
of the present invention is the uniformity of density of the semiconductor nanoparticles
within stratum 101. Density may range from 10^{12} to 10^{13} particles/cm² but other
densities may be used according to other exemplary embodiments. The stratum 101
is essentially a monolayer of discrete nanoparticles. A plurality of the discrete
15 nanoparticles, such as exemplary nanoparticles 111 and 113, contact each other. In
the preferred embodiment, the majority of adjacent nanoparticles may contact each
other. It is an advantage of the present invention that the individual nanoparticles that
are each capable of storing an electrical charge, are electrically isolated from one
another, even if in contact. This is attributable to the presence of the dielectric coating
20 formed on the particles prior to deposition on the substrate. Because of the ultra high
purity nitrogen used and the cleaning techniques used to maintain the processing
equipment in an ultra clean state, the density of foreign particle contamination on
surface 76 and therefore in stratum 101, due to the deposition of stratum 101, may be
less than 10^{11} atoms per square centimeter. The discrete nanoparticles of stratum 101
25 are characterized by a tightly controlled range of particle sizes. According to one
exemplary embodiment, the nanoparticle cores may be less than 15 nanometers.
According to an exemplary embodiment, the nanoparticles may be essentially the same
size. According to an exemplary embodiment, particle sizes of the nanoparticles of
stratum 101 include a particle size distribution that is approximately log-normal with
30 respect to diameter, with a geometric standard deviation less than 1.5. The log-normal
distribution is well-known in the art and is used to describe a population of samples that
is normally distributed with respect to the logarithm of the distribution parameter (e.g.,
diameter). In an exemplary embodiment, the nanoparticles may have an average core
diameter within the range of 2-10 nanometers or 3-7 nanometers and, also in an
35 exemplary embodiment, at least 90% of the cores will include a diameter within the
range of 3-7 nanometers. As noted above, cores 102, such as shown in Figure 6, are

1 preferably formed of single crystal material such that the semiconductor nanoparticle
may be referred to as a nanocrystal. According to yet another exemplary embodiment,
at least 90% of the cores of the discrete semiconductor nanoparticles may lie within a
5 range of 2-6 nanometers, or 90% of the cores may be less than 10 nanometers.

Figure 8 shows a cross-sectional view of a preferred embodiment of stratum 101.
As shown in Figure 8, the plurality of adjacent nanoparticles 100 contact each other
laterally. Because of oxide shells 106 formed on each nanoparticle 100, adjacent
semiconductor cores 102 are electrically insulated from one another and lateral shorting
10 is precluded.

In the subsequent description, semiconductor nanoparticles 100 will be referred
to as silicon nanoparticles, but the reader is reminded that such is exemplary only.
Semiconductor nanoparticles 100 may be formed of other semiconductor or metal
materials, alternatively. Figure 9 is a cross-sectional view showing stratum 101 of silicon
nanoparticles formed between subjacent oxide layer 153 and superjacent oxide layer
151. Stratum 101 is as shown and described in conjunction with Figures 7 and 8 and
includes silicon nanoparticles as shown and described in Figure 6. According to this
exemplary embodiment, cores 102 essentially function as a discontinuous layer of
silicon nanoparticles, each surrounded by an oxide including the oxide shells (present,
but not distinguishable in Figure 9) and the subjacent and superjacent oxide layers.
The oxide shell serves to electrically insulate silicon cores 102 from one another in a
monolayer formed within an oxide film formed over surface 155.

According to another exemplary embodiment, force manipulation of the
deposited silicon nanoparticles may be carried out. In an exemplary embodiment,
contact mode atomic force microscopy may be used to move nanocrystals that were
initially in spatially random positions such as deposited and as shown in Figure 7. An
ordered structure of nanocrystal "wires" and other ordered arrays may be produced
using this technique. Colloidal forces and self-assembly may alternatively be used to
produce ordered deposits.

According to a preferred embodiment, the stratum of silicon nanoparticles
including highly controlled particle sizes, densities and of superior uniformity may be
utilized as the floating gate of a non-volatile memory, floating gate transistor. According
to this embodiment, the stratum of silicon nanoparticles is formed over an oxide,
referred to as a tunnel oxide, formed over gate regions of transistors. Figure 10 shows
such an exemplary structure after the various layers have been formed and the layers
35 patterned to form a gate structure. In Figure 10, tunnel oxide film 124 is formed over

1 surface 76 of substrate 74 including over gate area 120. Stratum 101 of silicon
nanoparticles, including cores 102, is formed over tunnel oxide 124. Upper gate oxide
5 film 128 and gate electrode film 132 are next sequentially formed. After the sequence
of films are formed over one another, including over gate area 120, situated between
source/drain regions 122, conventional patterning and etching means are used to define
the gate region and to form gate electrode structure 140. A plurality of such structures
are preferably formed simultaneously according to the above process.

Gate structure 140 includes gate width 136, which may be less than 0.2 microns.
10 According to one exemplary embodiment, gate width 136 may range from 0.18 microns
to 1.2 microns, but other gate widths may be used according to other exemplary
embodiments. Thickness 126 of tunnel oxide film 124 and thickness 130 of upper gate
oxide film 128 are chosen in conjunction with the oxide shell thickness formed around
each of the silicon nanoparticles in stratum 101. The thicknesses are chosen such that
15 the total, or effective tunnel oxide thickness 127, lies within the range of 3-12
nanometers and in a preferred embodiment may lie within the range of 4-8 nanometers.
According to one exemplary embodiment, thickness 126 may be 3-12 nanometers, or
preferably 3-6 nanometers. Similarly, thickness 130 of upper gate oxide film 128 is
chosen so that total, or effective thickness 129 of the upper gate oxide lies within the
20 range of 6-15 nanometers. In an exemplary embodiment, thickness 130 or effective
thickness 129 may lie within the range of 10-15 nanometers, but other thicknesses
sufficient to preclude electrons exchanging between cores 102 and gate electrode 132
may be used alternatively. Various conventional films suitable for use as gate electrode
132 may be used and may be formed to various suitable thicknesses 134. Transistor
25 gate structure 140 is formed over channel region 120 and between source/drain regions
122. According to other exemplary embodiments, each of tunnel oxide film 124 and
upper gate oxide 128 may be formed of other dielectric materials. As discussed
previously, the discrete silicon nanoparticles of stratum 101 may be of various
dimensions and will preferably be crystallized silicon. Stratum 101 may alternatively be
30 described as a discontinuous floating gate of gate structure 140. Particle size may be
chosen in conjunction with gate width 136. In an exemplary embodiment, the total
diameter of a silicon nanoparticle 100, including oxide shell 106, will be chosen to be
less than $0.1 \times$ gate width 136. In this manner, at least ten silicon nanoparticles will be
included within stratum 101 across gate width 136. Additionally, the non-volatile
35 memory transistor formed to include gate structure 140 will include a length (depth into

and out of the plane of the drawing) ranging from 10-100 microns long according to an exemplary embodiment. Other transistor lengths may be used alternatively.

Now turning to Figure 11, floating gate transistor 162 is shown and includes gate structure 140. Floating gate transistor 162 is formed to include stratum 101 of aerosol silicon nanoparticles, and features program/erase characteristics comparable to conventional stacked gate non-volatile memory devices, excellent endurance (greater than 10^5 program/erase cycles), and long-term non-volatility, even when a thin bottom oxide 124 of less than 6 nanometers is used. The reader is reminded that stratum 101 may be formed of nanoparticles formed of other semiconductor or metal materials, in other exemplary embodiments. Floating gate transistor 162 includes source/drain contacts 154 formed through dielectric material 152, and may be electrically coupled, biased and tested through power supplies 150, 156, 158 and 160. Power supply 150 is used to bias gate electrode 132, power supplies 156 and 158 are connected to the source/drain regions 122, and power supply 160 is coupled to substrate 74 and channel region 120. Other conventional arrangements may be used for biasing exemplary transistor 162 and conventional means may be used for programming, erasing and testing transistor 162.

Typical subthreshold and output characteristics of exemplary aerosol-silicon nanocrystal field effect transistors are shown in Figures 12 and 13, respectively. Figure 12 is a graph of the drive current, I_d , versus gate voltage, V_{gate} , and Figure 13 is a graph of the drive current versus drain voltage, V_{DRAIN} . The values of the subthreshold slope (200mv/dec), and the DIBL (100mv/v) and the drive current (30 microamps/micron), are typical for thick gate dielectric, high substrate doped non-volatile memory devices. The threshold voltage, V_t , has been defined as the gate voltage corresponding to a drain-source current of 1 μA when a drain bias of 1 volt is applied. Uniform Fowler-Nordheim tunneling has been used for both the program and the erase operation, though programming using channel hot-electron injection is possible as well. Programming at source and drain separately also allows multi-bit storage in one device.

Figure 14 is a graph of threshold voltage versus time and shows programming transients for various gate voltages V_g , using a substrate bias of -5V. Figure 15 is a graph of threshold voltage versus time and shows erase transients for various V_g 's, using a substrate bias of -8V. As shown in Figures 14, 15, 16 and 17, the high areal nanocrystal-density obtained by the aerosol fabrication process results in a large threshold voltage window of at least two volts. The device programs to a high V_t of about +3.3 volts in 50 μs with gate and substrate bias of +8V and -5V, respectively.

1 The device erases to a low V_t of about +1V in 100ms with gate and substrate bias of
+8V and -8V, respectively. According to this exemplary embodiment, the threshold
voltage window is about 2.3 volts, but other threshold voltage windows may be
5 produced according to other exemplary embodiments.

Figure 16 is a graph indicative of endurance characteristics and showing the
program and erase V_t as a function of number of program/erase cycles. As shown in
Figure 16, the aerosol non-volatile memory devices feature excellent endurance
behavior as demonstrated by the small V_t window closure observed after 5×10^5
10 program/erase cycles. It can be seen that less than a 15% closure of the
program/erase V_t window is observable after 5×10^5 program/erase cycles. The gradual
shift of the V_t window to higher threshold voltages is indicative of charge being built up
in the gate oxide layer during cycling. Additionally, in spite of thin tunnel oxides used,
reasonable disturb times and long non-volatility is obtained indicative of intrinsic
15 advantages of nanocrystal charge storage. No stress induced leakage current has
been observed. No drain disturb was detected even using drain voltages as high as 4
volts, which is indicative of the absence of lateral conduction between the nanocrystals
in stratum 101.

Figure 17 is a graph showing retention and disturb characteristics (threshold
voltage - V_t , versus time) of an exemplary MOSFET formed to include a floating gate
20 formed of exemplary silicon nanocrystals of the present invention.

The preceding merely illustrates the principles of the invention. It will thus be
appreciated that those skilled in the art will be able to devise various arrangements
which, although not explicitly described or shown herein, embody the principles of the
invention and are included within its spirit and scope. For example, the present
25 invention can be used to provide various and multiple strata of silicon and other
semiconductor nanoparticles for various applications in the
semiconductor/microelectronics industry. According to various exemplary embodiments,
multiple strata of various nanoparticles may be formed directly on one another
according to the process of the present invention.

30 All examples and conditional language recited herein are principally intended
expressly to be only for pedagogical purposes and to aid the reader in understanding
the principals of the invention and the concepts contributed by the inventors to
furthering the art, and are to be construed as being without limitation to such specifically
35 recited examples and conditions. Moreover, all statements herein reciting principles,
aspects, and embodiments of the invention, as well as specific examples thereof, are

1 intended to encompass both structural and functional equivalents thereof. Additionally,
it is intended that such equivalents include both currently known equivalents such as
5 equivalents developed in the future, i.e., any elements developed that perform the same
function, regardless of structure. The scope of the present invention, therefore, is not
intended to be limited to the exemplary embodiments shown and described herein.
Rather, the scope and spirit of the present invention is embodied by the appended
claims.

10

15

20

25

30

35